

DOCUMENT RESUME

ED 227 157

TM 830 174

AUTHOR Chapline, Elaine Burns  
 TITLE Formative Evaluation in the Development of a Math Anxiety Reduction Program.  
 SPONS AGENCY Women's Educational Equity Act Program (ED), Washington, DC.  
 PUB DATE Apr 81  
 NOTE 17p.; Paper presented at the Annual Meeting of the American Educational Research Association (65th, Los Angeles, CA, April 13-17, 1981).  
 PUB TYPE Speeches/Conference Papers (150) -- Reports - Evaluative/Feasibility (142)  
 EDRS PRICE MF01/PC01 Plus Postage.  
 DESCRIPTORS Class Activities; \*Curriculum Development; Feedback; Females; \*Formative Evaluation; \*Mathematics Anxiety; Mathematics Curriculum; \*Mathematics Materials; Program Development; \*Sex Differences; Sex Stereotypes

ABSTRACT

The Teacher Education and Mathematics project, funded by the Women's Educational Equity Act, developed curriculum materials to increase mathematical knowledge and confidence, increase the perception of mathematics as a female domain, reduce math anxiety, and develop skills in identifying and counteracting sex bias in mathematics instruction. To develop effective materials, a formative evaluation was to document materials' use at each stage of development, to assess student responses to materials, to collect judgment of external experts, and to collect reactions and suggestions of instructors who used the materials. Evaluation issues explored in this paper are: reliance on observer descriptions and participant ratings, effects on students and instructional staff of intensive observations, effects on students of frequent requests for feedback and evaluative reactions, and the extent to which professionals can be expected to use materials as they have been developed. (Author/PN)

\*\*\*\*\*  
 \* Reproductions supplied by EDRS are the best that can be made \*  
 \* from the original document. \*  
 \*\*\*\*\*

ED227157

FORMATIVE EVALUATION IN THE DEVELOPMENT  
OF A MATH ANXIETY REDUCTION PROGRAM\*

Elaine Burns Chapline  
Queens College of  
The City University of New York

U.S. DEPARTMENT OF EDUCATION  
NATIONAL INSTITUTE OF EDUCATION  
EDUCATIONAL RESOURCES INFORMATION  
CENTER (ERIC)  
This document has been reproduced as  
received from the person or organization  
originating it  
Minor changes have been made to improve  
reproduction quality  
Points of view or opinions stated in this docu-  
ment do not necessarily represent official NIE  
position or policy

Mathematics is a key area in the educational curriculum. Mathematical competence is a necessity for persons who wish to function effectively in our increasingly technological society. Mathematics has been identified as fulfilling a gatekeeping or filtering function because of its importance in many careers (Sells, 1973; Burton, 1979). Sex differences in mathematical experiences both in and out of school, math anxiety and sex role stereotyping of mathematics as a male domain have frequently been cited as factors in math avoidance for women (Fennema and Sherman, 1977). Fox (1976) pointed out that

Although anxiety about mathematics may not in every case be a direct result of sex-role socialization conflicts, it is likely that the sex-typing of mathematics as a male domain by parents, teachers and peers results in the acceptance of math anxiety in females as inevitable or irrelevant to their development. (p. 45)

Math anxiety can begin as early as the elementary school years. Donady, Kogelman and Tobias (1976) found evidence of this phenomenon in women's reports during counseling sessions. Teachers' attitudes,

\*Paper presented at AERA Annual Meeting, Los Angeles, California, April 1981. The program described here has been developed under a grant from the U.S.D.E. Women's Educational Equity Act. The views expressed here are the responsibility of the grantee and no official endorsement by the U.S.D.E. should be inferred.

PERMISSION TO REPRODUCE THIS  
MATERIAL HAS BEEN GRANTED BY  
E. B. Chapline

TO THE EDUCATIONAL RESOURCES  
INFORMATION CENTER (ERIC).

TM 830 174



classroom activities, and sex role stereotyping in materials have all served to reinforce the perception of mathematics as a male domain (Brophy & Good, 1970; Fennema, 1980). The need for teachers who are free of math anxiety and therefore able to help their students to enjoy and to continue to study mathematics has been pointed out (Donady and Tobias, 1977). In order to foster students' mathematics skill development and positive attitudes, teachers need to perceive mathematics as equally important and useful for both girls and boys (Skypek, 1980).

Teacher Education and Mathematics (TEAM), a two-year project, funded by WEEA was designed to sensitize prospective elementary school teachers to the importance of mathematics for all students by developing a model program and instructional materials. These materials were designed to reduce teacher education students' mathematics anxiety; to increase their perception of mathematics as a female domain and to develop their skills in identifying sex bias in mathematics curriculum materials and teacher-pupil interaction in the classroom. (Within a teacher education program they can be used most advantageously before students take their course in mathematics curriculum and methods.) This paper deals with the formative evaluation strategies used during the TEAM project.

TEAM Program and Materials

Instructional strategies were designed by Dr. Claire Newman, Mathematics Educator and by the author, with the collaboration of the TEAM staff, both to foster positive attitudes toward math and to develop mathematical knowledge and skills through integrated cognitive and affective learning experiences. The inverse relationship between math anxiety and math knowledge reported in the literature (Aiken, 1970 and



3

1976; Richardson and Suinn, 1972 and 1973) provided the rationale for the program's stress on mathematical competence as well as confidence.

Four mathematics modules were developed, each of which consists of an Instructor's Text, Student Materials, Exercises, Solutions and Guide to Student Exercises, and Student Summary and Review. The Instructor's Text provides direction on what an instructor may do, suggests how instruction may be carried out and provides commentary which includes explanations, instructional options, attitudinal interventions and organizational ideas. The mathematics modules titles are: Patterns, Approximation and Estimation, Metric Measurement and Choice and Chance.

Modules designed to deal with students' attitudes consist of an Instructor's Text and Student Materials. Demystifying Mathematics deals with the individual's own math experiences, math myths which can serve as learning barriers, and math in daily life uses. Sex Role Stereotyping in Mathematics Education provides experiences in identifying sex bias in curriculum materials and classroom practices. Women, Mathematics and Careers provides opportunities for students to consider the role of mathematics in careers and to devise experiences for children to alert them to mathematics' role in careers. Women as Mathematicians introduces and describes women who have made significant contributions to mathematics.

An Instructor's Handbook outlines the structure and purposes of the instructional material. It includes: User's Guide, Introduction (to the problem of mathematics attitudes), Purposes and Strategies, Learning Environment, Attitudinal Interventions and Bibliographies.

TEAM materials were used first at Queens College in two sections of a mathematics course (N = 44) during one semester of the first year. In

4

2

the second year, some or all of the modules were used in courses in five settings. The fall classes were labelled TEAM II (N = 34) and the spring class TEAM III (N = 29).

#### Formative Evaluation Plan

Since TEAM was essentially a program and materials development project, it was necessary that the formative evaluation design assess the degree to which the products were, or were showing promise of, serving the project's goals. The formative evaluation focused on: documenting the uses of materials at each stage of their development, assessing students' responses to the materials, collecting judgments of materials from a panel of experts, and collecting reactions and suggestions from instructors who used the materials.

#### General Procedures

The evaluation/revision process was viewed as an interactive one, with evaluation staff interacting with program staff to clarify major points of emphasis in student objectives and teaching activities, and revising the evaluation data collection instruments to reflect these goals. Elenor Rubin Denker, who served as TEAM's Evaluation Associate provided feedback quickly to the staff on all data collected so that it could be considered in planning subsequent instructional sessions.

#### Instruments

Data gathering procedures were devised to serve specific project needs. In the first year of the project, during the initial materials development, a Rating Scale for Modules was devised which was used by consultants in their appraisal of the draft modules. This 13 item scale included ratings on the materials' usefulness, appropriateness for teacher education students, clarity, comprehensiveness, length, etc.

Ratings were made on a 5-point scale, ranging from "highly satisfactory" to "needs major revision." These ratings provided the basis for revisions and a second draft of the sample modules.

A Background Questionnaire which dealt with student's high school, and college mathematics course history, perceptions of family attitudes toward mathematics, ratings of mathematics achievement, attributions of selected mathematics course grades, and current feelings about studying mathematics was devised and piloted during the first semester of the project, and used in the second semester. Student responses to this questionnaire provided a profile of the teacher education students and finally of TEAM students which the instructors used in developing plans for specific learning experiences and class sessions.

During the second half of the first year, when the first group of students were enrolled in the TEAM course, an Initial Structured Interview was devised which contained a wide range of items of concern to the project, i.e. students' current feelings about mathematics, their mathematics learning history and details of any difficulties they had encountered and who had helped them with their difficulty, their perception of their family's interest in mathematics study, their anticipation of their feelings about studying mathematics currently and what they expected and wanted to get from the TEAM program. These interview data helped the staff to understand students' needs and starting points in the program. Until these data were available, specific needs of the TEAM students were unclear so projections and plans could be made only on the basis of the staff's experiences with former students.

An End of Semester Questionnaire, using a critical incident approach, asked what the best and worst things about the TEAM course were

and asked students to identify factors to which they would attribute a good and a poor math grade. An End of Semester Interview dealt with each student's description of her current math attitudes, current rating of math anxiety, changes (if any) in attitudes toward math, reactions to mathematical situations and any changes noted in the way a student noticed math in daily life, reanalysis of personal math history to share any new understandings developed about influences on attitudes, perceptions of teachers' influence on girls' mathematical self-confidence, feelings about the way females are portrayed in math texts, anticipated feelings about teaching math, feelings about being in an experimental course, and recommendations for the TEAM course. While these data fulfilled the function of assessing the effects of the program and materials on the first group of students, they served to provide feedback to the staff about which areas were satisfactory and which needed modification.

A Rating Scale for Class Sessions was developed and revised to provide ratings of math content (i.e. interesting, useful, new to you, understandable, frightening, hard) and ratings of the general atmosphere of the class (i.e. helpful, reassuring and enjoyable). Students' comments were also requested in the form. This scale was used at the end of the orientation session and after classes 3, 6, and 12.

During the second year or field trial stage, two questionnaires were developed in parallel form and used as a pre- and post-semester instrument to replace the interviews and questionnaires used in the first year. Since the materials were now to be used by teams other than the developers in settings in which the cost benefits of activities were

important for the new users, it was important to provide procedures that would be useful and not excessively costly in faculty time and effort.

The rating scale for classes which had been used in the first year was maintained as the instrument to gather students' responses to their class experience.

Another data source in year two were Logs in which students were asked to write a reaction each week, which emphasized their feelings about math. They were free to write about anything they wanted to discuss or needed to express. Student Logs were also kept during the TEAM III classes. Instructors stressed the nonevaluative nature of these Logs, that is they were not to be considered in assigning a student's grade for the course.

Instructors' reactions to the materials were sought through post session interviews by the person who was the assigned observer or by a staff member. These interviews included assessments of the materials used as well as suggestions for modifications.

Observations

During the first year or development stage, observational data were collected in class sessions by at least two observers. In general, we avoided the problem of inter-observer reliability by assigning one observer to a task. In some sessions there were observers, each with a specific task. All of these data were being gathered for the purpose of revising the materials before the field trials.

Participation data were tabulated, i.e. number of student speakers tallied during half-hour intervals, throughout the classes. These data were gathered because the goal of the project was to create conditions

in which students would actively participate. Verbal statements were used as an indicator of active participation.

Indications of anxiety were recorded by one observer who noted behavior such as giggling, restlessness, whispering, glancing around etc. and rated the degree of anxiety on a 4-point scale. Recognizing that these may also be signs of inattention due to factors other than anxiety, we decided to continue to gather these indicators and to note the content or topics periodically to see if these behaviors varied in relation to content.

A Log of each class session noted changes from the proposed math instructional sequence, times, students' questions, or areas of confusion, additional explanations provided, additional illustrations used, etc. Statements specifically directed to students' attitudes and the behavior or incidents that lead to them were noted along with the nature of the discussion that followed. During the second year of the project, this observation plan was reduced and simplified. One person kept a log of the flow of each session and noted the topic, activity and class structure and anxiety at 10 or 20 minute intervals.

#### Findings and Observations

While the Rating Scale for Modules yielded some useful data in that the degree to which a module was satisfactory on the specified criteria was indicated, the most useful information came in the comments and suggestions for modification. Discussions with consultants provided the richest source of information on module modification. The Rating Scale, and its criteria were useful, perhaps, in directing consultants' attention to features considered important by the developers.

Interview and Questionnaire data proved useful in describing the nature of the student group. In general, TEAM students were college juniors who had not taken mathematics courses since high school, who reported negative feelings and experiences about mathematics and who wanted to increase their achievement and confidence in working with mathematics. The students who were recruited for the classes were the ones the project had intended to serve. The variation found in age and ethnicity in the TEAM student group was considered desirable, since it had been the project's intent to develop materials that would be useful in diverse groups.

In general, the End of Semester Questionnaire and Interview data provided a picture of gains toward the project goals, e.g. 61% of the group rated themselves as less math anxious, 78% indicated that their general attitudes toward math had changed and were highly positive (30%) or somewhat positive (45%), 93% indicated awareness of sex stereotyping in curriculum materials, etc. Student's responses alerted the staff to the various ways in which students were conceptualizing and summarizing their experiences. Some people (39%) dealt with confidence in their ability to learn math, while others (21%) described the classroom atmosphere as supportive, and still others (16%) focused on their math achievement. Problem areas were identified through the variations in students' perceptions of the difficulty of the math content from one module to another, and in the appropriateness of the math content for some individuals. These reactions were considered in preparing the revisions of the modules for the second year.

The observation data dealing with the content and flow of the sessions proved particularly useful in module revision. Using these

data, some module sections were varied to provide more illustrative materials while others were reduced; more exercise materials that involved simpler applications of concepts were included, etc.

Participation data showed that an average of 75% of the students spoke at least once during each class session.

The observations of anxiety indicators turned up a particularly intriguing set of behaviors. Preening was noted during a highly anxious time when preparation for a quiz was being discussed. At that time, students were observed taking out hair brushes, nail files, lipstick, etc. and grooming themselves during the discussion. It would almost seem as if some of the women in the class were seeking reassurance that they would be acceptable in the face of the threatened loss of self-esteem posed by the quiz.

Both more frequent and longer students' ratings of the class sessions had been planned than actually occurred. Initially, each session was to have been rated, with a longer rating scale, but the staff settled for having each module rated with the revised, simplified scale described above. Students were willing to provide the ratings but they communicated in subtle ways that they were more interested in doing the class activities than in rating them. And too, the pressure on the evaluation staff to handle the quantity of data was a consideration in reducing the number of ratings. Ratings showed generally that students found the sessions satisfactory and found the general atmosphere supportive. The comments which students offered provided useful data for specific modifications that could be considered by the staff. Knowing the groups' averages in ratings provided a framework within which such individual suggestions could be considered.

During the second year, Logs provided a steady flow of information on students' feelings and reactions. This device proved especially useful as feedback for the class itself. For example, the tendency which the Logs revealed for students to explain their successes on quizzes as a function of the quiz's being easy rather than of their mastery of the materials was described to the class. An interesting discussion of the factors to which the students attributed their successes and failures followed. The Logs also gave the staff prompt feedback on the students' reactions to classes, assignments, mathematical situations and incidents, etc. Areas that needed staff attention as well as the degree to which students were facing and dealing with their anxieties was evident in their statements.

Issues and Suggestions

The complexity of telling the story of a case, of the descriptive narrative, has been detailed by Smith (1978). Trying to catch the essence of what was transpiring in the classes was our goal. One of the advantages of the design we used was the inclusion of groups over three semesters. This moved us beyond some of the limitations of the "one-shot case study" and it created problems in our knowing whether the changes in groups' responsiveness were a function of the improvements made in the materials and programs or of the differences between the groups.

Among the issues that became apparent during the use of the formative evaluation strategies described above was the heavy reliance of our data on observer's descriptions and participant and consultant ratings. These ratings were limited by the factors or criteria that had been identified for focus and the nature of the scales constructed for use.



While the criteria selected and items constructed were related logically to the project's objectives and included factors usually considered in evaluating curriculum materials, some uncertainty remains about the validity of the ratings. It should be noted, however, that while we were in the midst of the process, we tended to look upon our ratings and observational data as "real" indicators of the phenomena under consideration. "Project press," described by Smith (1977) was clearly a factor in making early decisions about the foci of the various observations and rating scales. Our limited resources gave rise to caution about committing heavily to an ethnographic model. We compromised and selected variables that seemed promising.

The presence of numbers of observers in the class sessions presents particular problems. Tensions were engendered in some of the participants despite detailed explanations that the purposes of the observation were to assess the materials and the instructor's procedures. While many of the students were able to forget, or ignore the observer's presence, others were not able to do so; they felt that they were in some way being studied or watched. It seems reasonable to conclude that a class is changed in some ways by the presence of an observation team. This issue has particular salience since the current increase of ethnographic classroom studies has brought and will continue to bring observers into classes. It may be that the psychological nature of some of the classes' content lead to students' increased reactivity to the observers' presence, and that the brief, intense nature of the project didn't permit for a kind of adaptation to the observers as "just part of the scene" for some people.

Participant attrition in TEAM was minimal. One factor which may have contributed to this was the students' having volunteered to participate. Having the program carried out in a class for which students were earning credits probably contributed to their commitment and served to maintain the group intact. Attrition is a factor of considerable consequence in experimental programs which have less structure.

The effects on students of frequent requests for feedback and evaluative reactions are unknown. It seems plausible that these requests increased students' self-consciousness. In some sense they became both the subjects and objects of study. While some students responded very well to this and seemed to grow in their ability to identify and express their feelings, others did not do so. For example, on Log entries, some students took a "Dear Diary" approach and were able to sense and describe what was happening to them. Others, in contrast, described the realities or external features of situations rather than their feelings or reactions. The Logs appear to yield useful data, as well as to provide an interesting growth experience for some people. Some students seemed amenable to providing descriptions and ratings while others were much more interested in spending their time in activities.

The extent to which professionals can be expected to implement a program and use materials developed by others in the recommended manner seems to depend upon many factors. The assumption underlying the establishment of field trials for materials is that they will be used in the ways in which the developers intended. It seems more reasonable to assume that strict adherence to a prescribed set of materials and procedures is not necessarily demanded--rather that instructors can

interact with the materials and use them effectively with a heightened sensitivity and awareness of students' needs. During the TEAM trials, instructors operated with the spirit along with many of the specifics of the program, that is they were modeling an approach which we want the students to take when they become teachers. It is also of interest to note that the levels of use of an innovation (Hall et al., 1975) place refinement, "the user varies the use of the innovation to increase the impact on clients within the immediate sphere of influence" at a higher point than routine use, ". . . few if any changes are being made in ongoing use" (p. 54). This emphasizes the value of materials being used to serve effectively the needs of particular students.

In summary, the TFAM formative evaluation design had as its purposes the documentation of use at each stage of materials' development, the assessment of student responses to the materials, the collection of judgmental responses from external experts and the collection of reactions and suggestions from instructors who used the materials. Issues raised by these strategies were identified.

### References

- Aiken, L.R. Attitudes toward mathematics. Review of Educational Research, 1970, 40, 551-596.
- Aiken, L.R. Update on attitudes and other affective variables in learning mathematics. Review of Educational Research, 1976, 46, 293-311.
- Brophy, J.E. and Good, T.L. Teachers, communication of differential expectation for children's classroom performance: Some behavioral data. Journal of Educational Psychology, 1970, 61, 365-374.
- Burton, G. Regardless of sex. Mathematics Teacher, 1979, 72, 261-270.
- Donady, B., Kogelman, S., and Tobias, S. Math anxiety and female mental health: Some unexpected links. Paper presented at the meeting of the National Advisory Committee on Mental Health Services and Women, Cambridge, Mass., August, 1976.
- Fennema, E. Teachers and sex bias in mathematics. The Mathematics Teacher, 1980, 73, 169-173.
- Fennema, E. and Sherman, J.A. Mathematics attitudes scales: Instruments designed to measure attitudes toward the learning of mathematics by females and males. Journal for Research in Mathematics Education, 1976, 7, 324-326.
- Fennema, E. and Sherman, J.A. Sexual stereotyping and mathematics learning, The Arithmetic Teacher, 1977, 24, 369-372.
- Fox, L.H. The effects of sex role socialization on mathematics participation and achievement. Paper presented at the Education and Work Group, Career Awareness Division, National Institute of Education, December, 1976.
- Hall, G.E. et al. Levels of use of the innovation: A framework for analyzing innovation adoption. Journal of Teacher Education, 1975, 26, 53-56.
- Richardson, F.C. and Swinn, R.M. The mathematics anxiety rating scale: Psychometric data. Journal of Counseling Psychology, 1972, 19, 551-554.
- Sells, L. High school mathematics as the critical filter in the job market. Unpublished paper, March, 1973.
- Smith, L.M. An evolving logic of participant observation, educational ethnography and other case studies. In Shulman, L.S. (Ed.) Review of Research in Education, Itasca, Illinois: F.E. Peacock Publishers Inc., 1978.

Smith, L.M. Science education in the Alte schools: A kind of case study. National Science Foundation Case Studies in Science Education. Washington, D.C. N.S.F., 1977.

Skypek, D.H. One point of view: Girls need mathematics, too, The Arithmetic Teacher, 1980, 27, 4-7.

Suinn, R.M. et al. The MARS, a measure of mathematics anxiety: Psychometric data. Journal of Clinical Psychology, 1972, 28, 373-375.